Duarte, C., Ferreira, R. J. D. O., Mackie, S. L., Kirwan, J. R., & Pereira da Silva, J. A. (2015). Outcome Measures in Polymyalgia Rheumatica: A Systematic Review. *Journal of Rheumatology*, *42*(12), 2503-2511. https://doi.org/10.3899/jrheum.150515

Title: A Systematic Review of Outcome Measures Used in Polymyalgia Rheumatica

Authors: Cátia Duarte, Ricardo J. O. Ferreira, Sarah Louise Mackie, John Richard Kirwan and José António Pereira da Silva, on behalf of the OMERACT Polymyalgia Rheumatica Special Interest Group

ABSTRACT

Objective: The objective of this work was to identify the instruments used to assess polymyalgia rheumatica (PMR) in published studies.

Methods: A systematic literature review of Clinical Trials and longitudinal observational studies related with PMR, published from 1970 to 2014, was carried out. All outcomes and assessment instruments used were extracted and categorised according to core areas and domains, as defined by OMERACT 11, filter 2.0.

Results: 35 articles (3,221 patients) were included: 12 RCTs; 3 non-randomised trials; and 20 observational studies. More than 20 domains were identified, measured by 29 different instruments. The most frequently used measures were pain, morning stiffness, patient and physician global assessment, erythrocyte sedimentation rate and C-reactive protein. The definition of outcomes varies considerably between studies.

Conclusion: The outcome measures and instruments used in PMR are numerous and diversely defined. The establishment of a core set of validated and standardized outcome measurements is needed.

Keywords: Polymyalgia Rheumatica, Outcome Assessment, Review Literature, OMERACT

Departments and Institutions

Centro Hospitalar e Universitário de Coimbra, EPE, Coimbra, Portugal

Faculdade de Medicina, Universidade de Coimbra, Coimbra, Portugal

Health Sciences Research Unit: Nursing (UICISA:E), Coimbra, Portugal

Leeds Institute of Rheumatic and Musculoskeletal Medicine and NIHR-Leeds Musculoskeletal Biomedical Research Unit, University of Leeds, Leeds, UK

Academic Rheumatology Unit, Bristol Royal Infirmary, Bristol, UK

Funding Sources: none

Affiliations

1 – Duarte C, MD, Rheumatology Department: Centro Hospitalar e Universitário de Coimbra, EPE, Coimbra, Portugal and Faculdade de Medicina, Universidade de Coimbra, Coimbra, Portugal

2 – Ferreira RJO, RN, MSc, Rheumatology Department: Centro Hospitalar e Universitário de Coimbra, EPE, Coimbra, Portugal and Health Sciences Research Unit: Nursing (UICISA:E), Coimbra, Portugal

3 – Mackie SL, BM BCh, PhD, Leeds Institute of Rheumatic and Musculoskeletal Medicine and NIHR-Leeds Musculoskeletal Biomedical Research Unit, University of Leeds, Leeds, UK

4 – Kirwan JR, PhD, University of Bristol, Academic Rheumatology Unit, Bristol Royal Infirmary, Bristol, UK

5- Da Silva JAP, PhD, Rheumatology Department: Centro Hospitalar e Universitário de Coimbra, EPE, Coimbra, Portugal and Faculdade de Medicina, Universidade de Coimbra, Portugal

e-mail Address of each author

Cátia Duarte: catiacmduarte@gmail.com

Ricardo J. O. Ferreira: ferreira.rjo@gmail.com

Sarah Louise Mackie: S.L.Mackie@leeds.ac.uk

John Richard Kirwan: John.Kirwan@bristol.ac.uk

José António Pereira da Silva: jdasilva@ci.uc.pt

Corresponding author (reprints and correspondence)

Dr.ª Cátia Duarte

Serviço de Reumatologia, Piso 7

Centro Hospitalar Universitário de Coimbra, EPE

Avenida Dr. Bissaya Barreto,

3000-075 Coimbra. Portugal

catiacmduarte@gmail.com

Telephone: (00351) 239 400 547 Fax: (00351) 239 400 587

Short running footline: Outcome Measures in PMR

INTRODUCTION

PMR is an inflammatory disease with a lifetime risk estimated at 2.4% for women and 1.7% for men (1) and a peak incidence after 60 years of age. The diagnosis of PMR relies on clinical and laboratory manifestations, supported by a rapid, favourable response to glucocorticoid therapy at medium doses (15-20 mg/day of prednisone or equivalent). When untreated, PMR can cause profound disability. Glucocorticoid therapy remains the gold standard therapy for PMR and is usually efficacious. However, the potential toxicity of long-term glucocorticoid therapy (2) imposes the need to search for safer alternatives.

Future research in PMR requires the use of valid and reliable outcome measures that encompass the relevant scope of disease manifestations. A variety of outcomes have been used to assess disease activity, including clinical features (pain and morning stiffness), ultrasonography parameters and laboratory measures such as Erythrocyte Sedimentation Rate (ESR), C-Reactive Protein (CRP) and interleukin-6 (IL-6) levels. Composite scores of disease activity (3) and definitions of good response, remission and relapse have been proposed (3-6). However, these measures have not yet been extensively validated in PMR and do not incorporate patient viewpoints.

The Outcome Measures in Rheumatology (OMERACT) initiative aims to develop core sets of outcome measures capable of providing consistent estimates of the benefits of interventions for each given condition in clinical trials (7). According to the OMERACT 11 filter 2.0, such core sets should include at least one Domain from each Core Area. Four Core Areas, broad aspects of a health condition, are defined: three encompass the "Impact of Health conditions" – *Life Impact*, *Resource Use*, and *Death*, and a fourth Core Area encompasses *Pathophysiological Manifestations* (8, 9). This filter also considers Domains, as sub-specifications within one Area (9, 10). In order to be included in a core set, a domain should be measurable by truthful, discriminative and feasible instruments (9, 11).

The OMERACT PMR Working Group was formed to define a core set of outcome measures to be used in future clinical research in PMR. With this systematic literature review we aim to supply this endeavour with objective information on the outcome measures currently used to assess PMR disease activity and response to treatment.

METHODS

Search strategy: The literature search was performed in MEDLINE, CINAHL, Science Citation Index of Web of Sciences, Cochrane Library (Cochrane Central Register of Controlled Trials CENTRAL and Cochrane Database of Systematic Reviews CDSR). The research strategy was based on the following Key words: ("Polymyalgia Rheumatica"[MeSH]), and covered material published from January 1st 1970 to June 30th 2014.

Inclusion Criteria: Studies were included if they: 1) used published classification criteria to select patients; 2) were written in English, French, Portuguese or Spanish languages; 3) followed a design of either Clinical Trial or longitudinal observational study, and 4) were available in full text. Studies which included an heterogeneous patient sample and published data that did not allow differentiating PMR subjects from other diseases (eg. Giant Cell Arteritis or late onset Rheumatoid Arthritis) were excluded.

Study selection: Titles, abstracts and full reports of articles identified were systematically and independently screened by two authors (CD and RF) with regards to inclusion and exclusion criteria. In the first step, selection was based on titles and abstracts. Full reports of articles selected on this phase were evaluated (second step) to select the papers to include in this systematic review. Disagreements regarding selection of one article were discussed between both reviewers until consensus was reached. Persistent disagreements were resolved by a third evaluator (JAPS).

Data extraction: During data extraction, special attention was given to the "Patients and Methods" and "Results" sections of each article. All data were extracted using a standardized

template designed for this review, which had been piloted and improved, which included: study design, sample size, follow-up period, outcome measures used and the method of assessment.

Each outcome was characterized according with OMERACT Filter 2.0 considering Core Areas (Pathophysiological Manifestations, Life Impact, Death, Resource Use) and Domains (10).

RESULTS

The results of the literature search and selection of papers is presented in Figure 1. The electronic search strategy yielded 868 articles, 43 of which were selected, on the basis of title and abstract, for further assessment/detailed review. At the end, 35 studies (12-46) met the inclusion criteria for this systematic review (Figure 1). The agreement between the two reviewers was 96.6 % and 100% on the first and second step of articles' selection, respectively.

(Place Figure 1 here)

Included studies

Table 1 shows the study design characteristics of the included articles. Twelve of the included studies are RCTs of medication, controlled against either placebo or conventional PMR treatment (12-23). Three are non-randomized interventional studies or without clear information about randomization (24-26). Longitudinal observational studies represent more than half of the selected articles (20 of 35) (27-46). One of these observational studies (36) is a long-term follow-up of one RCT already included (20). The study size ranged from 4 (24) to 781 subjects (32, 41), with follow-up periods varying between 14 days (22) to 34 years (32). All studies included a majority of females and patients older than 50 years, which is in agreement with classical PMR features (47-50).

(Place Table 1 here)

The studies identified in this literature review include outcomes and instruments pertinent to all Core Areas defined by OMERACT, except Resource Use. The Core Area most represented is "Pathophysiological Manifestations" which included a total of six domains, followed by "Life Impact" with five domains (Table 2).

(Place table 2 here)

Pain

Pain was used as an outcome in 17 studies (12-16, 18, 19, 22, 25, 34, 36, 39, 40, 43-46) A Visual Analogue Scale (VAS) was commonly used to quantify pain, usually as a 0-10 cm scale (11 studies). In three of them (12, 14, 16), pain was graded using an ordinal scale form "0" to "3".

Most published reports do not provide a clear definition of the pain being assessed. The description of pain localization varies: "shoulder and pelvic girdle pain" (15), "proximal pain" (34), "proximal muscle pain" (14, 16), or "joint or muscle pain" (13). Matteson and colleagues evaluated pain considering different locations as shoulder, limbs and global (44). None of the published reports specified the period of time under evaluation when asking the patients about their "pain".

Morning Stiffness

Stiffness, more commonly morning stiffness, was considered in almost all the included studies (13-16, 18-22, 24-28, 30, 33-38, 40, 43-46). It was evaluated as an independent outcome in 11 studies (13-16, 18, 19, 25, 27, 34, 40, 44), and was included as parameter in composite disease activity scores or in the definition of relapse/recurrence/remission in an additional 15 studies (19, 20, 22, 24, 26, 28, 33, 35-38, 43-46).

Morning stiffness was measured in terms of duration ("*minutes"*) in the majority of the studies. In one RCT (18), morning stiffness duration was reported in one study through a 4-point scale (1= <30 min.; 2= 30-60 min.; 3= 60-120 min.; and 4= >120 min.) In two studies, stiffness was graded from 0 to 3, where "0" means no symptoms, but it is unclear whether severity, duration or both were being assessed (14, 16). No information is given to the meaning of the other values in the scale. Only Weyand and colleagues (27) evaluated the severity of morning stiffness using a 0-10 cm VAS. Only one RCT (13) and one observational study (27) gave precise information about the time interval under evaluation ("*average of last week*").

Patient and Physician Global assessment

Patient Global Assessment (PGA) of disease activity was measured in nine studies (13, 19, 22, 25, 27, 33, 35, 38, 46), always as a 0-10 cm VAS except in two studies (13, 27), where a 5-point ordinal scale was used.

Physician Global Assessment (PhGA) was used in 14 studies (12, 13, 19, 22, 25, 27, 33, 35, 36, 38, 43-46), twelve of them as a 0-10 cm VAS and two (13, 27) as 5-point ordinal scales.

In 9 studies (22, 25, 35, 36, 38, 43-46) both PGA and PhGA were included as a parameter within a pre-defined composite disease activity score.

Two instruments were employed by a single study (33): 1. a PGA of General Health and 2. a Patient Satisfaction with Disease Status (PATSAT) (classification of disease state according to the Austrian school mark system - 1=excellent, 2=good, 3=average, 4=moderate, 5=unsatisfactory).

Function and Quality of Life (QoL)

Function was assessed in five observational studies (34, 36, 38, 44, 46), one open label trial (25) and three RCTs (21-23). In all studies, function was assessed through the generic instrument Health Assessment Questionnaire (HAQ) (51).

Health-related QoL was considered in two large observational studies (34, 44) and was assessed through the generic tool Medical Outcome Survey-Short Form (MOS-SF 36) (52). In a single observational study (46), QoL during the past month was assessed using a 0-100 mm VAS, where 0 means normal and 100 the worst QoL.

Other Clinical outcomes: Elevation of upper limbs was considered as an outcome in some studies (22, 25, 33, 35, 43, 45, 46), always as a component of a composite disease activity score. Upper limb elevation was measured on a 0-3 scale with the following levels: 3 - no upper limb elevation; 2 - elevation bellow shoulder level (<90º); 1 - elevation at shoulder level (90º); and 0 - elevation above shoulder level (>90º). **Muscle function** (23) (hand grip strength and jump test), chair stand test (23), ten meters walking (23), and **time to onset of fatigue** (hours) (13) were used as outcomes in one single study each. Intensity of fatigue reported by the patient, using a 0-100 mm VAS, was assessed in a single study (44).

ESR and CRP

ESR (12-19, 21, 22, 24-28, 30, 31, 33-40, 43-46) and CRP (12, 13, 15, 17, 19, 21, 23-26, 28, 30, 31, 33-38, 40, 43-46) were used in the assessment of disease activity by most but not all RCT and observational studies.

Other Laboratory measures

Other laboratory outcome measures used in some observational and clinical trials include serum fibrinogen (12, 15, 16, 45, 46) and IL-6 levels (19, 22, 24, 27, 30, 31, 37), mainly as experimental evaluations.

Ultrasonography

Ultrasonography was used in three prospective observational studies (38, 40, 44) and in one open label trial (25). Different studies used different evaluation protocols, there being no formal proposal for the standardization of US evaluation of response to therapy in PMR. Jiménez-Palop and colleagues considered the evaluation of intra-articular synovitis at the shoulder and hip, tenosynovitis, and bursitis in the shoulder. This study demonstrated good inter and intra-observer reliability (0.96 and 0.99 respectively) but no statistically significant correlation was found with clinical and laboratory parameters of disease activity (40).

Composite Measures

Most of the studies integrated the individual outcome measures into composite indices, considered as response/relapse criteria or activity scores. This is summarized in table 3. Most of them defined relapse or recurrence as the observation of new symptoms, increase of ESR (usually >30 mm), or increase of CRP (higher than 0.5 mg/dl or 1 mg/dl), after remission has been achieved, in patients receiving glucocorticoids (GCs) or after discontinuation of GCs, respectively. Proposed Response Criteria include an improvement of symptoms and reduction/normalization of inflammatory parameters (ESR and CRP). In 2003, the European Collaborating Polymyalgia Rheumatica Group proposed a core set of response criteria. These EULAR response criteria comprise an improvement in VAS pain (obligatory) and at least 3 of the following 4 items: CRP (mg/l) or ESR (mm/1st h), Morning stiffness (min), Elevation of upper limbs (0–3), and VAS PhGA (4). However, there is considerable discrepancy in the definition of "improvement" and in the duration of improvement required to define "response".

One of the most common composite disease activity scores used was the Polymyalgia Rheumatica Activity Score (PMR-AS), developed by Leeb and Bird (6) and defined as

PMR-AS = CRP(mg/dl) + **E**(0-3) + 0.1X**MST**(min) + **VASp**(0-10) + **VASph**(0-10)[1]

The PMR-AS score, showed a good correlation with other outcome measures, namely with VAS Patient Global assessment (r=0.76) and ESR (r=0.32) (6, 33). Given that CRP is a component of PMR-AS it is not surprising that the composite score correlated with ESR, which is closely associated with CRP. Similarly, another component of PMR-AS is the patient pain VAS, and patient global VAS is usually strongly correlated with pain VAS. PMR-AS also showed a very good internal consistency in two different cohorts (Cronbach alfa of 0.90 and 0.88) (6) and demonstrated reliability (3, 33, 53).

(Place Table 3 here)

Glucocorticoid Therapy

The characterization of the GC treatment regime employed is extremely variable. Only a few studies included the cumulative dose of GCs (21, 27, 36), the minimum dose required (13, 17, 21), the duration of therapy (21), or the percentage of discontinuation of steroids after a specified duration of follow up (13, 20, 21, 36).

Adverse Events

The incidence and characterization of adverse events related to interventions were described in the majority of the clinical trials (15-25) and in the long-term follow-up study of patients treated with methotrexate (36). None of the studies performed a systematic and structured evaluation of safety.

Some observational studies were designed to assess specific adverse events related to GCs, such as vertebral fractures (39, 42), bone mineral content (16, 42), cardiovascular and cerebrovascular events (32, 42). One study described mortality and its causes (29). The

[1] CRP: C-Reactive Protein; E: Elevation of upper limbs; MST: Morning Stiffness Time; VAS: Visual Analogue Scale; p: patient; ph: physician.

methods used to elicit adverse effects in observational studies was variable but death registries and patient files were the most common sources of information.

DISCUSSION

This systematic literature review highlights a remarkable variability in the assessment of PMR in research settings.

Patient Reported Outcomes (PRO) are the most commonly studied outcomes and were assessed in almost all of the studies included in this review. Fatigue, however, was evaluated in two studies only. Function and QoL were evaluated in less than 10% of the studies, in spite of the importance to patients of these (54).

The instruments used to measure PROs in the selected articles were very heterogeneous. Also, there was, in general, a poor definition of what is actually being measured (e.g. concerning morning stiffness: is the question referring to the girdles, the hands or elsewhere? At what time of the day? What is the time period being assessed? Are we measuring duration, severity or both?).

There are no studies addressing the relative importance of each outcome from the patients' perspective. During the "OMERACT 11" Meeting (North Carolina, USA, May 2012) data from a preliminary, "scoping" consultation exercise involving 104 PMR patients from three centres from the UK and one from Belgium was presented by the PMR-SIG Group. In this study, patients were invited to express their concerns regarding disease and treatment. Symptoms and "impairment" were clearly important to patients, with pain, stiffness, fatigue, and sleep disturbance being mentioned very often. Physical activity and treatment aspects like glucocorticoid-related adverse effects were also considered important (54). It is important that patients' concerns and wishes are incorporated into any core outcome set.

Outcomes assessed by physicians rather than patients were less heterogeneous. Physician reported outcomes were used less frequently than PRO, with PhGA (as a 0-10cm VAS) being the most commonly used. Given the discrepancies between patient' and physicians' evaluations that have been found in several diseases (55-57), it is generally considered that both PRO and physician reported outcomes should be included to capture the burden of disease.

All selected articles reported either ESR or CRP, except in studies designed to evaluate specific adverse events (29, 32, 41, 42). Other laboratory parameters, such as IL-6 and fibrinogen, or ultrasonography have been considered so far as "experimental" outcomes.

Disease activity scores or definitions of Remission, incorporating both physician- and patient-reported outcomes, are well-established in other Rheumatic Diseases and may prove useful also in PMR. The concepts of remission/relapse/recurrence are not consistently defined for PMR. A composite score of disease activity, the PMR-AS, was developed by Leeb and colleagues in 2007 (33) and has been used in approximately 40% of the selected articles published after 2007.

We recognize some strengths and limitations to our study. We used the most important databases of medical research articles, considered other languages beside English and scrutinized a long period of time. The lack of evaluation of the quality of papers may be seen as a limitation but we believe this was the most adequate strategy to serve the primary goal of identifying all possible outcomes under current use. As a limitation, we have not searched Conference abstracts or contacted the authors in order to enlarge our scope. By including only longitudinal observational studies and clinical trials with PMR population, we may have lost some outcomes used in cross-sectional studies or in larger studies of rheumatic diseases. We did not perform any psychometric analysis of each instrument, as this was outside the intended scope of this work.

CONCLUSION

Our study revealed that a great heterogeneity currently exists in the assessment of PMR. Most instruments have been insufficiently validated according to the OMERACT Filter and the patients' perspective may not always have been fully covered. These data suggest that further work is needed to define and validate relevant outcome measures for assessment of PMR in order to promote clinical research in this field and enhance comparability of studies. Core areas and domains need to be defined according with the OMERACT procedures. Evaluation instruments capable of satisfying the properties required by OMERACT Filter 2.0 need to be developed, including validity, reliability, feasibility, and responsiveness.

REFERENCES

1. Crowson CS, Matteson EL, Myasoedova E, Michet CJ, Ernste FC, Warrington KJ, et al. The lifetime risk of adult-onset rheumatoid arthritis and other inflammatory autoimmune rheumatic diseases. Arthritis Rheum. 2011;63(3):633-9.
2. Hoes JN, Jacobs JW, Boers M, Boumpas D, Buttgereit F, Caeyers N, et al. EULAR evidence-based recommendations on the management of systemic glucocorticoid therapy in rheumatic diseases. Ann Rheum Dis. 2007;66(12):1560-7.
3. Binard A, De Bandt M, Berthelot JM, Saraux A. Usefulness of the disease activity scores for polymyalgia rheumatica for predicting glucocorticoid dose changes: a study of 243 scenarios. Arthritis Rheum. 2007;57(3):481-6.
4. Leeb BF, Bird HA, Nesher G, Andel I, Hueber W, Logar D, et al. EULAR response criteria for polymyalgia rheumatica: results of an initiative of the European Collaborating Polymyalgia Rheumatica Group (subcommittee of ESCISIT). Ann Rheum Dis. 2003;62(12):1189-94.
5. Dejaco C, Duftner C, Cimmino MA, Dasgupta B, Salvarani C, Crowson CS, et al. Definition of remission and relapse in polymyalgia rheumatica: data from a literature search compared with a Delphi-based expert consensus. Ann Rheum Dis. 2011;70(3):447-53.
6. Leeb BF, Bird HA. A disease activity score for polymyalgia rheumatica. Ann Rheum Dis. 2004;63(10):1279-83.
7. Kirwan JR, Boers M, Tugwell P. Updating the OMERACT filter at OMERACT 11. J Rheumatol. 2014;41(5):975-7.
8. Kirwan JR, Boers M, Hewlett S, Beaton D, Bingham CO, 3rd, Choy E, et al. Updating the OMERACT filter: core areas as a basis for defining core outcome sets. J Rheumatol. 2014;41(5):994-9.
9. Boers M, Kirwan John R, Turgwell P, Beaton D, III COB, Conaghan PG, et al. The OMERACT Handbook: OMERACT; 2014 JULY 2014. 124 p.
10. Boers M, Kirwan JR, Wells G, Beaton D, Gossec L, d'Agostino MA, et al. Developing core outcome measurement sets for clinical trials: OMERACT filter 2.0. J Clin Epidemiol. 2014;67(7):745-53.

11. Boers M, Kirwan JR, Gossec L, Conaghan PG, D'Agostino MA, Bingham CO, et al. How to choose core outcome measurement sets for clinical trials: OMERACT 11 approves filter 2.0. J Rheumatol. 2014;41(5):1025-30.
12. Lund B, Egsmose C, Jorgensen S, Krogsgaard MR. Establishment of the relative antiinflammatory potency of deflazacort and prednisone in polymyalgia rheumatica. Calcif Tissue Int. 1987;41(6):316-20.
13. Littman BH, Bjarnason D, Bryant G, Engelbrecht J, Cohen M, Mertz L, et al. Steroid sparing activity of tenidap in patients with polymyalgia rheumatica: a multicenter double blind randomized placebo controlled study. J Rheumatol. 1995;22(6):1097-103.
14. Krogsgaard MR, Lund B, Johnsson B. A longterm prospective study of the equipotency between deflazacort and prednisolone in the treatment of patients with polymyalgia rheumatica. J Rheumatol. 1995;22(9):1660-2.
15. Di Munno O, Imbimbo B, Mazzantini M, Milani S, Occhipinti G, Pasero G. Deflazacort versus methylprednisolone in polymyalgia rheumatica: clinical equivalence and relative antiinflammatory potency of different treatment regimens. J Rheumatol. 1995;22(8):1492-8.
16. Krogsgaard MR, Thamsborg G, Lund B. Changes in bone mass during low dose corticosteroid treatment in patients with polymyalgia rheumatica: a double blind, prospective comparison between prednisolone and deflazacort. Ann Rheum Dis. 1996;55(2):143-6.
17. Ferraccioli G, Salaffi F, De Vita S, Casatta L, Bartoli E. Methotrexate in polymyalgia rheumatica: preliminary results of an open, randomized study. J Rheumatol. 1996;23(4):624-8.
18. Dasgupta B, Dolan AL, Panayi GS, Fernandes L. An initially double-blind controlled 96 week trial of depot methylprednisolone against oral prednisolone in the treatment of polymyalgia rheumatica. Br J Rheumatol. 1998;37(2):189-95.
19. Salvarani C, Cantini F, Olivieri I, Barozzi L, Macchioni L, Boiardi L, et al. Corticosteroid injections in polymyalgia rheumatica: a double-blind, prospective, randomized, placebo controlled study. J Rheumatol. 2000;27(6):1470-6.
20. Caporali R, Cimmino MA, Ferraccioli G, Gerli R, Klersy C, Salvarani C, et al. Prednisone plus methotrexate for polymyalgia rheumatica: a randomized, double-blind, placebo-controlled trial. Ann Intern Med. 2004;141(7):493-500.
21. Salvarani C, Macchioni P, Manzini C, Paolazzi G, Trotta A, Manganelli P, et al. Infliximab plus prednisone or placebo plus prednisone for the initial treatment of polymyalgia rheumatica: a randomized trial. Ann Intern Med. 2007;146(9):631-9.

22. Kreiner F, Galbo H. Effect of etanercept in polymyalgia rheumatica: a randomized controlled trial. Arthritis Res Ther. 2010;12(5):R176.
23. Bjorkman MP, Pilvi TK, Kekkonen RA, Korpela R, Tilvis RS. Similar effects of leucine rich and regular dairy products on muscle mass and functions of older polymyalgia rheumatica patients: a randomized crossover trial. J Nutr Health Aging. 2011;15(6):462-7.
24. Salvarani C, Cantini F, Niccoli L, Catanoso MG, Macchioni P, Pulsatelli L, et al. Treatment of refractory polymyalgia rheumatica with infliximab: a pilot study. J Rheumatol. 2003;30(4):760-3.
25. Catanoso MG, Macchioni P, Boiardi L, Pipitone N, Salvarani C. Treatment of refractory polymyalgia rheumatica with etanercept: an open pilot study. Arthritis Rheum. 2007;57(8):1514-9.
26. Cimmino MA, Parodi M, Montecucco C, Caporali R. The correct prednisone starting dose in polymyalgia rheumatica is related to body weight but not to disease severity. BMC Musculoskelet Disord. 2011;12(1):94.
27. Weyand CM, Fulbright JW, Evans JM, Hunder GG, Goronzy JJ. Corticosteroid requirements in polymyalgia rheumatica. Arch Intern Med. 1999;159(6):577-84.
28. Cantini F, Salvarani C, Olivieri I, Macchioni L, Ranzi A, Niccoli L, et al. Erythrocyte sedimentation rate and C-reactive protein in the evaluation of disease activity and severity in polymyalgia rheumatica: a prospective follow-up study. Semin Arthritis Rheum. 2000;30(1):17-24.
29. Myklebust G, Wilsgaard T, Jacobsen BK, Gran JT. Causes of death in polymyalgia rheumatica. A prospective longitudinal study of 315 cases and matched population controls. Scand J Rheumatol. 2003;32(1):38-41.
30. Salvarani C, Cantini F, Niccoli L, Macchioni P, Consonni D, Bajocchi G, et al. Acute-phase reactants and the risk of relapse/recurrence in polymyalgia rheumatica: a prospective followup study. Arthritis Rheum. 2005;53(1):33-8.
31. Boiardi L, Casali B, Farnetti E, Pipitone N, Nicoli D, Cantini F, et al. Relationship between interleukin 6 promoter polymorphism at position -174, IL-6 serum levels, and the risk of relapse/recurrence in polymyalgia rheumatica. J Rheumatol. 2006;33(4):703-8.
32. Kremers HM, Reinalda MS, Crowson CS, Davis JM, 3rd, Hunder GG, Gabriel SE. Glucocorticoids and cardiovascular and cerebrovascular events in polymyalgia rheumatica. Arthritis Rheum. 2007;57(2):279-86.

33. Leeb BF, Rintelen B, Sautner J, Fassl C, Bird HA. The polymyalgia rheumatica activity score in daily use: proposal for a definition of remission. Arthritis Rheum. 2007;57(5):810-5.
34. Hutchings A, Hollywood J, Lamping DL, Pease CT, Chakravarty K, Silverman B, et al. Clinical outcomes, quality of life, and diagnostic uncertainty in the first year of polymyalgia rheumatica. Arthritis Rheum. 2007;57(5):803-9.
35. Binard A, de Bandt M, Berthelot JM, Saraux A. Performance of the polymyalgia rheumatica activity score for diagnosing disease flares. Arthritis Rheum. 2008;59(2):263-9.
36. Cimmino MA, Salvarani C, Macchioni P, Gerli R, Bartoloni Bocci E, Montecucco C, et al. Long-term follow-up of polymyalgia rheumatica patients treated with methotrexate and steroids. Clin Exp Rheumatol. 2008;26(3):395-400.
37. Pulsatelli L, Boiardi L, Pignotti E, Dolzani P, Silvestri T, Macchioni P, et al. Serum interleukin-6 receptor in polymyalgia rheumatica: a potential marker of relapse/recurrence risk. Arthritis Rheum. 2008;59(8):1147-54.
38. Macchioni P, Catanoso MG, Pipitone N, Boiardi L, Salvarani C. Longitudinal examination with shoulder ultrasound of patients with polymyalgia rheumatica. Rheumatology (Oxford). 2009;48(12):1566-9.
39. Calvo L, Pistone G, Arnone S, Colomba D, Amico S, Giacalone A, et al. Polymyalgia rheumatica and vertebral fractures: a 1-year pilot controlled study. Rheumatol Int. 2010;30(9):1245-7.
40. Jiménez-Palop M, Naredo E, Humbrado L, Medina J, Uson J, Francisco F, et al. Ultrasonographic monitoring of response to therapy in polymyalgia rheumatica. Ann Rheum Dis. 2010;69(5):879-82.
41. Kang JH, Sheu JJ, Lin HC. Polymyalgia rheumatica and the risk of stroke: a three-year follow-up study. Cerebrovasc Dis. 2011;32(5):497-503.
42. Mazzantini M, Torre C, Miccoli M, Baggiani A, Talarico R, Bombardieri S, et al. Adverse events during longterm low-dose glucocorticoid treatment of polymyalgia rheumatica: a retrospective study. J Rheumatol. 2012;39(3):552-7.
43. Cleuziou C, Binard A, De Bandt M, Berthelot JM, Saraux A. Contribution of the polymyalgia rheumatica activity score to glucocorticoid dosage adjustment in everyday practice. J Rheumatol. 2012;39(2):310-3.
44. Matteson EL, Maradit-Kremers H, Cimmino MA, Schmidt WA, Schirmer M, Salvarani C, et al. Patient-reported outcomes in polymyalgia rheumatica. J Rheumatol. 2012;39(4):795-803.

45. McCarthy EM, MacMullan PA, Al-Mudhaffer S, Madigan A, Donnelly S, McCarthy CJ, et al. Plasma fibrinogen is an accurate marker of disease activity in patients with polymyalgia rheumatica. Rheumatology (Oxford). 2013;52(3):465-71.
46. McCarthy EM, MacMullan PA, Al-Mudhaffer S, Madigan A, Donnelly S, McCarthy CJ, et al. Plasma fibrinogen along with patient-reported outcome measures enhances management of polymyalgia rheumatica: a prospective study. J Rheumatol. 2014;41(5):931-7.
47. Chuang TY, Hunder GG, Ilstrup DM, Kurland LT. Polymyalgia rheumatica: a 10-year epidemiologic and clinical study. Ann Intern Med. 1982;97(5):672-80.
48. Bird HA, Esselinckx W, Dixon AS, Mowat AG, Wood PH. An evaluation of criteria for polymyalgia rheumatica. Ann Rheum Dis. 1979;38(5):434-9.
49. Healey LA. Long-term follow-up of polymyalgia rheumatica: evidence for synovitis. Semin Arthritis Rheum. 1984;13(4):322-8.
50. Jones JG, Hazleman BL. Prognosis and management of polymyalgia rheumatica. Ann Rheum Dis. 1981;40(1):1-5.
51. Wolfe F. A reappraisal of HAQ disability in rheumatoid arthritis. Arthritis Rheum. 2000;43(12):2751-61.
52. Jenkinson C, Stewart-Brown S, Petersen S, Paice C. Assessment of the SF-36 version 2 in the United Kingdom. J Epidemiol Community Health. 1999;53(1):46-50.
53. Binard A, Lefebvre B, De Bandt M, Berthelot JM, Saraux A. Validity of the polymyalgia rheumatica activity score in primary care practice. Ann Rheum Dis. 2009;68(4):541-5.
54. Mackie SL, Arat S, da Silva J, Duarte C, Halliday S, Hughes R, et al. Polymyalgia Rheumatica (PMR) Special Interest Group at OMERACT 11: outcomes of importance for patients with PMR. J Rheumatol. 2014;41(4):819-23.
55. Kwoh CK, O'Connor GT, Regan-Smith MG, Olmstead EM, Brown LA, Burnett JB, et al. Concordance between clinician and patient assessment of physical and mental health status. J Rheumatol. 1992;19(7):1031-7.
56. Neville C, Clarke AE, Joseph L, Belisle P, Ferland D, Fortin PR. Learning from discordance in patient and physician global assessments of systemic lupus erythematosus disease activity. J Rheumatol. 2000;27(3):675-9.
57. Hidding A, van Santen M, De Klerk E, Gielen X, Boers M, Geenen R, et al. Comparison between self-report measures and clinical observations of functional disability in ankylosing spondylitis, rheumatoid arthritis and fibromyalgia. J Rheumatol. 1994;21(5):818-23.

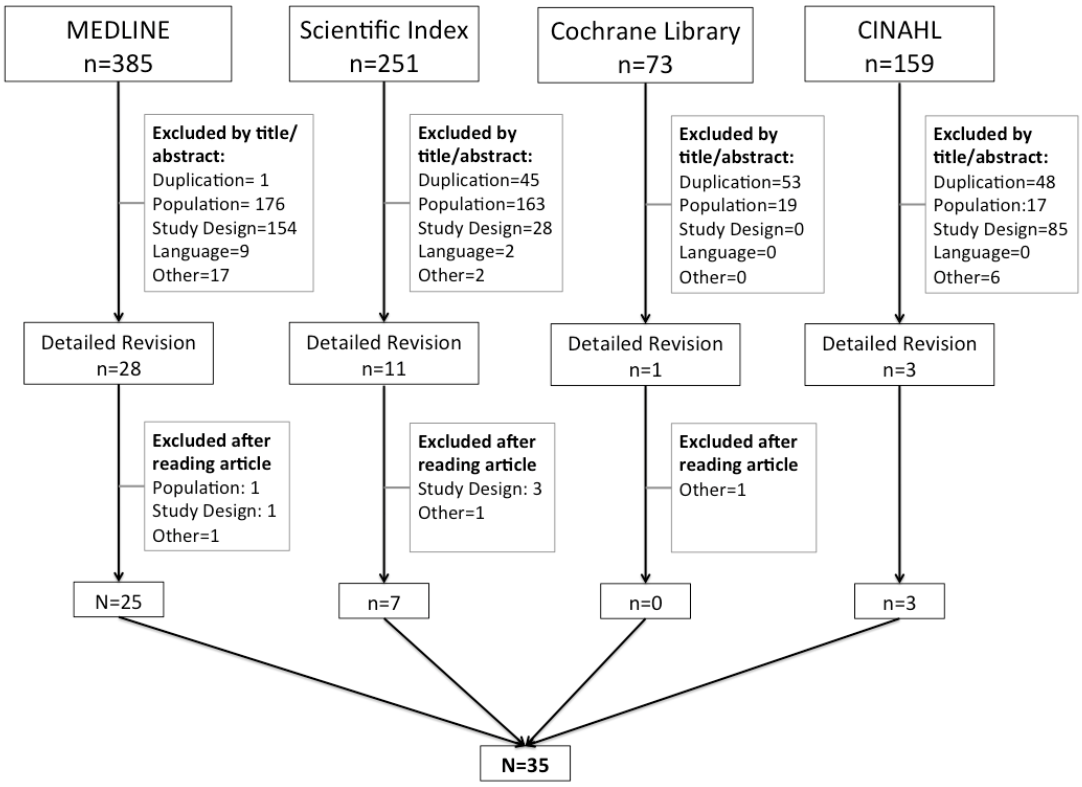

Figure 1: **Flow chart of the search and selection process.**

year	First Author and reference	Study design/ Protocol	Intervention	Sample size	PMR Definition/ stage of disease	Follow-up
Randomized clinical trials						
1987	Lund (12)	RCT, double blind, cross over (after a single blind parallel 3 arms)	Deflazacort vs Prednisone, weigh ratios of 1:1,2; 1: 1,5 and 1:1,8	41	Bird Maintenance phase	12 weeks
1995	Littman (13)	RCT, multicentre, placebo controlled	Tenidap 120 mg/day+ PDN 10 mg/day vs Placebo+PDN 10mg/Day	32	Study Definition Stable disease	15 weeks
1995	Krogsgaard (14)	RCT, double blind, controlled	Deflazacort vs prednisolone	30	Bird New diagnoses	12 months
1995	Di Munno (15)	RCT, double blind, cross-over	Deflazacort vs methylprednisone	29	Study definition New diagnoses	12 weeks
1996	Krogsgaard (16)	RCT, double blind, controlled	Deflazacort vs prednisolone	30	Bird New diagnoses	12 months
1996	Ferraccioli (17)	RCT, Open.	Prednisone 15 mg vs Prednisone 25mg+ Metothrexate 10mg/wk	24	Study definition New diagnoses	12 months
1998	Dasgupta (18)	RCT, double blind, multicentre	Methylprednisone depot VS prednisolone po	60	Study definition New diagnoses	96 weeks
2000	Salvarani (19)	RCT, double blind placebo controlled	Shoulder injection of 40 mg of methylprednisone	20	Healey New diagnoses	7 months
2004	Caporalli (20)	RCT, multicentre, double-blind, placebo controlled	Methotrexate 10 mg/wk+ GC vs Placebo+ GC	72	Chuang New diagnoses	18 months
2007	Salvarani (21)	RCT, multicentre, double blind, placebo controlled	Infliximab 3mg/kg vs Placebo	51	Healey New diagnoses	52 weeks
2010	Kreiner (22)	RCT, double blind, placebo controlled.	Etanercept 25 mg 2/wk vs Placebo	20	Chuang, New diagnoses	2 weeks
2011	Björman (23)	RCT, cross-over, double blind	Casein vs protein-enriched dairy product	60	Rheumatologist definition*	20 weeks
Non-randomized clinical trials						
2003	Salvarani (24)	Open, pilot, uncontrolled study	Infliximab + Prednisone	4	Healey Longstanding disease	49 weeks
2007	Catanoso (25)	Clinical trial, open, uncontrolled	Etanercept 25mg twice/wk, 24 wk	6	Healey Relapsing/ longstanding	36 weeks

2011	Cimmino (26)	Clinical trial not randomized, uncontrolled	Prednisone 12,5mg/id	60	Bird New diagnoses	6 months
Observational studies						
1999	Weyand (27)	Prospective observational study	NA	30	Study definition New diagnoses	12-33 months
2000	Cantini (28)	Prospective observational study	NA	177	Healey New diagnoses	5 years
2003	Myklebust (29)	Prospective observational study Gender-age matched controls	NA	65	Bird or Harlim Any stage	1987-1997
2005	Salvarani (30)	Prospective observational study	NA	94	Healey New diagnoses	24 months
2006	Boiardi (31)	Prospective observational study	NA	112	Healey New diagnoses	24 months
2007	Kremers (32)	Prospective observational study	NA	364	Chuand and Hunder, New diagnoses	1970-2004
2007	Leeb (33)	Prospective observational study	NA	39	Bird	18 months
2007	Hutchings (34)	Prospective observational, multicentre study	NA	129	Hazelman New diagnoses	12 months
2008	Binard (35)	Prospective observational study	NA	89	Rheumatologist definition*	Not defined
2008	Cimmino (36)	Long term follow-up of RCT (20)	Methotrexate 10 mg/wk	57	Chuang New diagnoses	5 years
2008	Pulsatelli (37)	Prospective observational study	NA	93	Healey New diagnoses	24 months
2009	Macchioni (38)	Prospective observational study	NA	57	Healy New diagnoses	41 months
2010	Calvo (39)	Case Cohort study	NA	20	ACR criteria (Chuang and Healey) New diagnoses	12 months
2010	Jiménez-Palop (40)	Prospective observational, multicentre study	NA	59	Study definition New diagnoses	12 weeks
2011	Kang (41)	Prospective	NA	781	Rheumatologist	3 years

		observational study			definition* New onset cases	
2012	Mazzantini (42)	Retrospective Observational study	NA	222	Bird Longstanding disease	Not defined
2012	Cleuziou (43)	Prospective observational study	NA	89	Rheumatologist definition*	Not defined
2012	Matteson (44)	Prospective observational study	NA	85	ACR/EULAR New diagnoses	6 months
2013	McCarthy (45)	Prospective observational study	NA	60	Jones & Hazleman New diagnoses	6 weeks
2014	McCarthy (46)	Prospective observational study	NA	60	Jones & Hazleman New diagnoses	6 weeks

***Table 1*: Characterization of the studies included in the analysis.**

NA: Not Applicable; RCT: Randomized Controlled Trial.* autonomous clinical diagnoses by attending rheumatologist.

Core Area	Domain	N. of studies	Instrument	N. of studies using the instrument
Pathophysiological	Laboratory markers	30	ESR	29
			CRP	23
			IL6, Fibrinogen	12
	Ultrasonography	4	Girdles US evaluation	4
	Pain	17	VAS 0-10 cm	11
			VAS 0-100 mm	2
			VAS 0-32	1
			Grade 0-3	3
	Morning stiffness	26	Duration (min)	11
			Grade	4
			Severity (VAS 0-10)	1
			As a parameter of *compositum* measure or definition	15
Life impact	PGA	9	VAS 0-10	7
			5 point-scale	2
	PhGA	14	VAS 0-10	12
			5 point scale	2
	PATSAT	1	Range 1-5	1
	Function	5	HAQ	5
	Quality of life	2	MOS-SF36	1
			VAS 0-100	1
Death	Mortality	1	SMR	1
Resource	-	0	-	-
Composite measures	Disease activity	8	PMR-AS	8
	Remission	6	Own definition	6
	Recurrence/Relapse	8	Own definition	8
Contextual factors Adverse events	Side effects	14	General	12
			Bone mineral content	2
			Vertebral fracture	2
	Glucocorticoid therapy	6	Minimal dose	2
			Cumulative dose	3
			Discontinuation	4
	Vascular disease	3	AMI, HF, CVA, PVD	3

Table 2: **Health Areas, Domains and Instruments reported in the 35 selected articles.** AMI: Acute Myocardial Infarction, CVA: Cerebrovascular Accident, HAQ: Health Assessment Questionnaire; HF: Heart Failure, PMR-AS: Polymyalgia Rheumatica Activity Score; PVD: Peripheral Vascular Disease, SMR: Standardized Mortality Rate; VAS: Visual Analogue Scale.

Year	Study	Good Response	Remission	Relapse (under GC´s) or Recurrence (after GC´s)	Activity Score
1995	Di Munno (15)	At the end of 2 weeks of GC´s: >50 % of pain, morning stiffness, ESR and CRP improvement At the end of 12 weeks of GC´s: >80 % of improvement in pain and morning stiffness; ESR and CRP normal	NA	NA	NA
2000	Cantini (28)	NA	NA	Joint signs or symptoms ESR>30 mm/hr Restart or increase GCs	NA
2000	Salvarani (19)	>70% improvement in Pain-VAS, Patient and Physical Global Assessment, and morning stiffness	NA	NA	NA
2003	Salvarani (24)	NA	NA	Typical symptoms Morning stiffness >1hour ESR >30 mm/hr CRP >0,5 mg/dl	NA
2004	Caporali (20)	NA	NA	Joint signs or symptoms (aching and stiffness of shoulder, hip girdle or both) ESR>30 mm/hr CRP>0,5 mg/dl	NA
2005	Salvarani (30)	NA	NA	Increase of symptoms ESR>30 mm/hr or CRP >0,5 mg/dl Good response after increase or restart GCs	NA
2007	Catanoso (25)	EULAR Response Criteria	NA	NA	PMR-AS
2007	Hucthings (34)	No pain or improvement of >50% in VAS girdles-pain, Morning stiffness<30 min, ESR <30 and CRP<1 mg/dl	NA	New symptoms and worsening lab tests requiring increase of GCs	

2007	Leeb (33)	NA	NA		PMR-AS
2007	Salvarani (21)	NA	No symptoms or signs Normal ESR	Increase of symptoms (aching and stiffness of shoulder, hip girdle or both) ESR>30 mm/hr or CRP >0,5 mg/dl Good response after increase or restart GCs	NA
2008	Binard (35)	NA	NA	NA	PMR-AS
2008	Cimmino (36)	NA	NA	Joint signs or symptoms (aching and stiffness of shoulder, hip girdle or both) ESR>30 mm/hr CRP >0,5 mg/dl	NA
2008	Pulsateli (37)	NA	NA	Recurrence of symptoms ESR>30 mm/hr, CRP>0,5 mg/dl, Good response after restart or increase GCs	NA
2009	Machioni (38)	NA	NA	Reappearance of clinical manifestations ESR >30mm/hr CRP>0,5 mg/dl	PMR-AS
2010	Kreiner (22)	NA	NA	NA	PMR-AS
2011	Cimmino (26)	NA	≥70% improvement in clinical symptoms of PMR ESR and CRP normal 1 month after start therapy	NA	NA
2012	Cleuziou (43)	NA	NA	NA	PMR-AS
2013	McCarthy (45)	NA	NA	NA	PMR-AS
2014	McCarthy (46)	NA	NA	NA	PMR-AS

Table 3: **Summary definitions of Good Response, Remission, Relapse, Recurrence and Disease Activity used in different studies.**

ESR: CRP: C-Reactive Protein; Erythrocyte Sedimentation Rate; GCs: Glucocorticoids; NA: Not Applicable/Available; PMR-AS: Polymyalgia Rheumatica Activity Score; VAS: Visual Analogue Scale. Relapse (increase of symptoms/signs after remission or good response, in patients still receiving GCs) or recurrence (reappearance of symptoms and lab changes after remission or good response, following discontinuation of GCs).